DOCUMENT RESUME

ED 209 251 TM 810 672

AUTHOR Roecks, Alan L.; Andrews, John H.
TITLE Levels of Use Interviews: A Successful Formative Evaluation Tool [Revised].
INSTITUTION Education Service Center Region 20, San Antonio, Tex.
PUB DATE Apr 80
NOTE 38p.; Paper presented at the Annual Meeting of the American Educational Research Association (64th, Boston, MA, April, 1980).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Elementary School Teachers; *Evaluation Methods; *Formative Evaluation; Inservice Teacher Education; *Intermediate Administrative Units; *Interviews; Program Costs; *Program Evaluation; Program Improvement; Reading Programs
IDENTIFIERS *Levels of Use of the Innovation

ABSTRACT

Levels of Use (LOU) interviews can be used for formative evaluation purposes in an intermediate education agency. Programs and services for training teachers provided to districts are evaluated. LOU interviews give program staff unique information for improving programs. Program improvement is aided by LOU interviews because information is provided that program staff cannot obtain on their own. The information corroborates whether programs are proceeding according to plan. Evaluation and program staffs develop a stronger working relationship and both gain an understanding of the program. Well defined relevant criteria is the goal of the process. Survey instruments do not reveal the same kind of information. The interview technique allows communication and effectively identifies how teachers are using the innovation. A wealth of formative evaluation information is provided by the interviews in the areas of student characteristics, materials, objectives, testing, and classroom management. The information proved useful in identifying areas of need and allowed project staff to make decisions for program improvement. It is especially valuable in its application to the purposes of formative evaluation by practitioners responsible for effective change in schools. The major limitation cited was the cost to execute valid evaluation. (DWH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

HI!! I'M LOUELLA, OR SIMPLY
LOU. I'M HERE TO TELL YOU
ABOUT
LEVELS OF USE INTERVIEWS:
A SUCCESSFUL FORMATIVE EVALUATION TOOL
L o U

Levels of Use Interviews: A Successful Formative Evaluation Tool

Alan L. Roecks
John H. Andrews

Evaluation Services
Education Service Center, Region 20
San Antonio, Texas

Since the passage of the Elementary and Secondary Education Act of 1965, evaluation[1] of educational programs often has relied on the experimental model. Answers were sought to the question "Does the program work or not?" All too frequently, studies resulted in "no significant findings" and programs were judged failures--despite experimental evidence suggesting that the programs did, indeed, produce desireable results. One explanation for this discouraging finding could be the failure of proper program implementation. Evaluating such programs runs the risk of appraising "non-events" (Charters & Jones, 1973). Research from the extensive Rand study on Title I ESEA programs (Berman & McLaughlin, 1978) revealed that the program in many districts had not taken place, suggesting that failure of proper implementation may be more widespread than previously thought. Freeman (1977) contends that many "no effect" evaluations can be explained by the failure of proper program implementation:

> There may be programs that offer significant solutions to defects in the human condition that have been passed over because evaluation failed to show an impact. Yet, in fact, the real reason for lack of impact was that the program was never implemented fully well, or at all. (p. 27)

Some researchers, including Mushkin (1973) and Rossi, Freeman and Wright (1979) see the problem due to faulty research procedures. Comparative

---

[1] The definition of evaluation used for this paper is taken from the Phi Delta Kappa's National Study Committee on Evaluation: "Educational evaluation is the process of delineating, obtaining, and providing useful information for judging decision alternatives." (Stufflebeam, Foley, Gephart. Guba, Hammond, Merriman, and Provus, 1971, p. 40).

educational studies may be based on instructional models not sensitive to the actual environment. The dominance of the scientific method limits serious consideration of alternative research paradigms. Variables that can be controlled are included in experimental studies; other relevant variables may be excluded. Writing a decade earlier, Stufflebeam et al (1971) wisely advised that evaluation methodology must not be limited to the experimental model:

> ...Perhaps the greatest challenge facing the evaluator is overcoming the idea that evaluation methodology is identical to research methodology. Equating them forces certain constraints inimical to the purposes of evaluation and makes it impossible to meet certain of the needs served by good evaluation. (p. 22)

Elaborating, the authors caution that conventional experimental design may not be appropriate in situations where the efficiency of the overall process, including desired outcomes, is studied:

> Perhaps the most damaging assertion about the application of conventional experimental design to evaluation situations is that it conflicts with the principle that evaluation should facilitate the continuous improvement of a program. Experimental designs prevent rather than promote changes in the treatments...
>
> It is probably unrealistic to expect directors of innovative projects to accept these conditions, because they obviously cannot constrain a treatment to its original, undoubtedly imperfect form, just to insure internally valid end-of-year data....concepts of evaluation are needed which could stimulate, not stiffle, dynamic development of programs. (p. 26)

Qualitative Methodology

Viable alternatives to the experimental model recently have been developed. Qualitative methodology has gained increasing support from evaluation investigators (Gebhardt, 1979). Because they provide insight into the fundamental processes of education, qualitative methodologies emphasize getting "closer to the data," developing an understanding of the observed and describing the reality of the situation (Rist, 1977). Common methodologies include participant or non-participant observation and comprehensive interviews.

Qualitative inquiry can be independent or complimentary of qualitative study. Qualitative procedures document the need for quantitative studies. Making sure proper program implementation has occurred, determined from qualitative inquiry, is often necessary before carrying out a more quantitatively-oriented impact study. Information from qualitative studies is useful in addressing the formative evaluation question "How can this program be improved?"

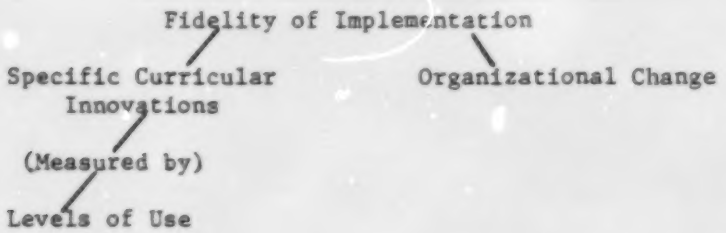
> ...If the agency is interested in information on its effectiveness in delivering services, the evaluation can study the process of program implementation and find out the extent to which the program is producing the quantity, quality, and coverage of services that were expected. (Weiss, 1972, p. 75).

The flexibility of qualitative studies allows formative evaluation of educational training programs, where findings can be used for the purpose of program improvement. Summative evaluation, which involves quantitative procedures, is harder to carry out due to the inherent difficulty in defining how training programs are implemented. Educational training programs often do not follow a prescribed curriculum; the goals of the program may be implemented differently across organizational levels (Anderson & Ball, 1978, Chapter 10). This paper describes how a qualitative methodology, Levels of Use (LoU), was applied successfully in the formative evaluation of a teacher training program.

Levels of Use Interviews

In a comprehensive review of curriculum and instruction implementation, Fullan and Pomfret (1977) revealed that a majority of research studies centered on the fidelity of implementation. Fidelity, which compares actual and intended use for the innovation, was researched in two distinct ways--focusing either on organizational change or specific curricular innovations. One technique identified by Fullan and Pomfret for measuring curriculum implementation

was Levels of Use (LoU) of the Innovation. The relationship of LoU to the concept of fidelity of implementation is illustrated below.



Conceptually, LoU is part of the Concerns Based Adoption Model (CBAM). CBAM accounts for implementation using two approaches. One approach focuses on changes in the innovation itself, the second identifies two critical dimensions for describing change from the individual perspective. These dimensions are Stages of Concern about the Innovation (Hall & Rutherford, 1976) and Levels of Use of the Innovation (Hall, Loucks, Rutherford, & Newlove, 1975). A series of longitudinal and cross-sectional studies have verified the existence of seven Stages of Concern and eight Levels of Use (See Appendix A). LoU has received limited use as a curricular evaluation tool with emphasis placed on summative evaluation. Okpalobi (1979) demonstrated the effectiveness of a reading staff development program by researching the overall and eight component LoU ratings for 27 treatment and 21 control teachers. With a few exceptions with some ambiguity at the higher end of the scale, Hall and Loucks (1977) and Reidy and Hord (1979) discovered a positive relationship between the LoU and student achievement. Hall and Loucks also found that teachers in their first year of use do not implement as well and correspondingly do not have as high LoU rating. Other facts about LoU include:

- Reliability. Interater reliability ranges from .87 to .96 for the overall LoU rating.
- LoU was developed by educators for educators.

- Validity. Ethnographic studies of all day teacher activity reveal a correlation of .98 between interview ratings and ratings from observation.
- Overall implementation. LoU can determine the degree of implementation for individuals exposed to treatment.

To the best of the authors' knowledge, LoU interviews have not been used for formative evaluation purposes. This paper documents how LoU can be used for this purpose in an intermediate education agency.

Region 20

The state of Texas is divided into twenty regions served by intermediate education service agencies. Established as part of the public education system of Texas in 1967, the service agencies offer school districts and teachers an opportunity to receive specialized services that normally would be beyond the reach of the average sized school system. By participating in programs with the intermediate agencies, school systems take advantage of cost-saving benefits and program financing and planning which result from cooperative efforts. They also realize additional funding and services that are available from state and federal services for regional programs of this type. Since Texas Service Centers are not in a regulatory position for the districts they serve, student achievement data generally is not available for evaluation of Center programs.

Evaluation at Region 20 has a dual thrust. First, technical assistance is provided to districts. Evaluation workshops are provided to district staff on topics including how Levels of Use can document whether or not the program

is in place. Second, programs and services for training teachers provided to districts are evaluated. Center staff represent evaluation "clients." Evaluation focuses on how teachers implement Center programs. LoU is one of several evaluation techniques employed.

The READS Program

READS was one of about 30 curricular programs housed at Region 20 during the 1978-79 school year. READS consisted of a sequence of student objectives and pre-posttests for grades 1-8. Teachers used these materials along with a record keeping system to individualize reading instruction. Funded from Title IV-C, ESEA monies, READs was was in its third (Phase III) and final year of operation when evaluated. The teacher sample reported in this paper consisted of teachers trained during the previous two years of the program.

The evaluation focus was both summative and formative. Objectives whose intent have been quantified describe program direction. The audience for summative evaluation, where a judgement about program worth is made, were the funding agency and Region 20 administration. The summative objective for READS required that at least 60% of the teachers trained during the two previous years be implementers.[2]

> By May 30, 1979, 60% of the teachers in Phase I and Phase II schools and 40% of the teachers in Phase III schools will be implementing READS with all or a section of their student population. Accomplishment will be shown by a representative sample of teachers judged implementers (LoU III, IV-A, IV-B, V).

The primary users of evaluation information are program staff. Their needs are met by evaluation information which address immediate, short term

---

[2]The major product of the evaluation office's first year (1977-78) was to develop measurable objectives for each Region 20 program. (See Drezek, Estes, Roecks, & Andrews, 1980.) A decision to use LoU interviews for measuring curricular implementation had not been made when the objectives for READS were submitted to the funding agency. Subsequently, objectives were modified so as to be measured by LoU.

needs. The formative evaluation objective for READS reflects this intent:

> By June 30, 1979, a revised edition of the English version of READS will be developed by project staff. Revisions will be based, in part, on information from LoU interviews with project teachers. Accomplishment of this objective will be evidenced by a panel of three reading curriculum experts selected by the ESC-20 evaluation component judging the revision to be meaningful and substantial.

Limitations

READS was one of several complicated programs at Region 20 during the late 1970s. Evaluating such an involved program proved difficult and time consuming. Moreover, the evaluation represented the first attempt at using LoU as a formative evaluation tool. For these reasons, the authors were hesitant to use information from READS for this paper. Despite these concerns, the evaluation of READS best shows how LoU interviews can be used in an applied setting. Evaluation findings collected were some of the most diverse and comprehensive available.

The next section outlines the procedures currently used to carry out LoU interviews to evaluate 14 curricular programs. These procedures are similar to the ones used for READS. Problems encountered in the pilot run with READS are illustrative of difficulties which may be encountered by others using LoU for the first time. They include:

- The original sampling scheme had to be discarded. About one-half of the teachers from the previous year could not be located.

- The list of participants was not complete. The actual number of teachers trained was not known.
- Criteria for implementation were not piloted. Five criteria were originally identified. One of the criteria was shown later to be critical for implementation.
- Insufficient information was collected for certain criteria. Especially difficult was collecting information for non-users. Subsequent data analysis was difficult.
- Interviews were done by five evaluators, although 70% of the interviews were collected by the same two evaluators preparing the final report.
- Given the above difficulties, the evaluation cost was nearly twice as much as budgeted.

LoU PROCEDURES

This section illustrates current procedures for carrying out an evaluation study using LoU. The five steps involved in this procedure are illustrated by examples from READS. Figure 1, presented at this time as an advanced organizer, gives the relative amount of evaluation time needed to carry out each step of the process. About two-thirds of an evaluator's available[3] time is divided evenly between conducting interviews and analyzing data. Another one-fifth is spent selecting and contacting the sample. The remainder of this section describes procedures associated with each step.

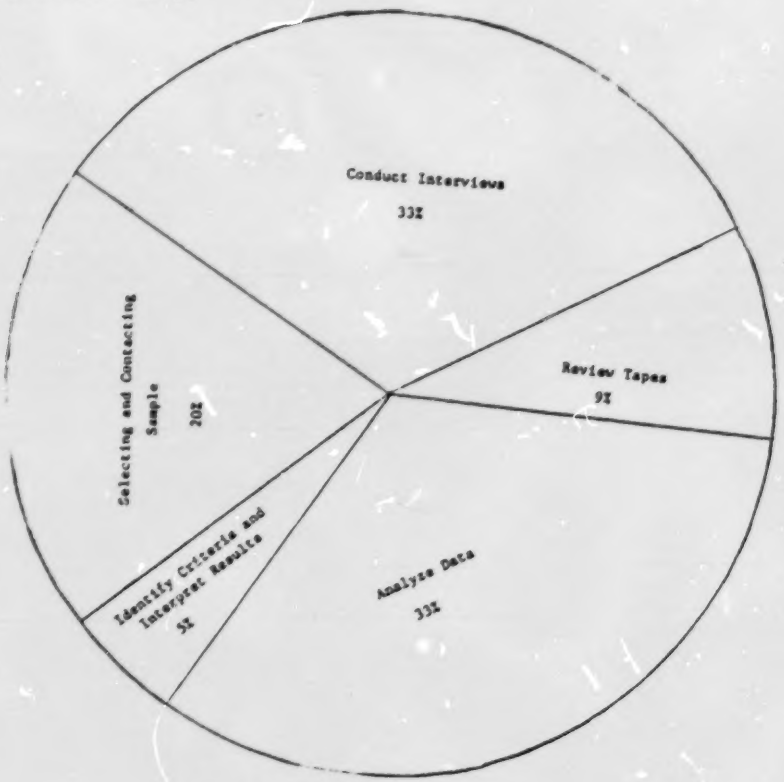


Figure 1
How an Evaluator's Available Time is Used

[3]Between 75% and 80% of an evaluator's time is available for direct evaluation work as documented by time and effort records. Remaining time is spent attending staff meetings, professional development activities and carrying out institutional responsibilities.

Identifying Implementation Criteria

A key role for evaluation is identifying the goals and objectives of a program (Rossi, 1978). Agreeing on what should be the criteria for implementation is a time-consuming and often frustrating process.

Securing agreement requires strong working relationship between evaluation and program staff.

In our setting, this relationship is often present due to the emphasis on formative evaluation where evaluation and program staffs work closely together. One characteristic unique to intermediate education agencies makes negotiating the criteria difficult.

The program developer and program facilitator are often the same person, titled the program manager.

The dual roles are in conflict. As program developer, the program manager wants to set the criteria strictly so that the description of the program can be put in its best light. However, in their role as change facilitator, the same program manager wants to have very general criteria to increase the probability that a higher percentage of teachers will be implementers.

Implementation criteria are piloted during the first year of a program[4]. When determining tentative criteria, evaluation and program staff must decide what the innovation "looks like" in the classroom. They must respond to such questions as:

> If teacher Smith participated in the program and is using the materials provided, and if teacher Jones does not participate at all, what would be happening to teacher Smith's class that wouldn't be happening to teacher Jones' class?

The criteria are then piloted on a few teachers. Face to face interviews are carried out presently with the possibility of using telephone interviews being explored. Pilot interviews reveal teachers who use programs in unusual ways or who are borderline with respect to one or more of the criteria. In reviewing these cases evaluation and program staff have real data, not hypothetical situations, which can be used to better define the boundaries of the program. From this, the criteria can be refined and finalized. New criteria may also be identified.

A clear definition of the innovation is essential in obtaining qualitative information useful to program staff, irrespective of the data collection instrument employed. Our experiences suggested two types of implementation criteria could be identified; coincidentally, the same conclusion was drawn by researchers

---

[4] The first year of operation centers on developing materials and ironing out wrinkles in the instructional system. Teachers may not be trained until the program has been operationalized six months or longer. Implementation is seen as the second and third year priority. Criteria identified during the first year remains essentially the same for succeeding years.

of the University of Texas Research and Development Center for Teacher education (See Hall & Loucks, 1978). These criteria are called "critical" and "related." Teachers must meet critical criteria in order to be implementing the program. Critical criteria represent primary program goals identified after the program is operational. Related criteria, which also describes individual implementation, are seen by the program manager as important but not essential for implementation. Teachers who are users meet the implementation criteria and most related criteria. The finalized criteria for READS are given below:

Critical Criteria:

1. The teachers must be using the student objectives contained in the program. They might skip a few objectives or use them in a slightly different order but, in general, they should follow the prescribed sequence.

Related Criteria:

1. The teacher must identify the reading level (skills) of each individual child.
2. The students must be grouped by skill levels (one student can be considered a group).
3. The teacher must be testing for skill mastery by objective. The mastery level can be established by the teacher.
4. The teacher must have a beginning at organizing teaching resources to go with the objectives. This will be a retrieval system such as card file or file folder.

Identifying the Sample and Scheduling Interviews

The sampling design depends on program size (how many districts, campuses, and teachers are served) and complexity (time of implementation is proportionate to complexity). The number of participant interviews is proportionate to project level of funding. Sampling is stratified randomly by campus. About twenty teachers are interviewed per project. The relatively small sample size means that data is aggregated for all teachers and is not analyzed by district or campus level.

Identifying the teacher sample can be a frustrating task. Teacher turnover can be as high as 50%. This problem is acute when trying to locate previous participants in the program. Finding teachers who have changed campuses can also be perplexing. The sample drawn must be increased 30% to 50% of required size. Replacement must be taken into account.

A second problem is that the list of participants, maintained by program staff, may be inaccurate. Nothing is more perplexing than calling a teacher who says "I am not involved in the program and haven't seen the program staff for over one year!"

Scheduling interviews in our setting is an involved process. Time seems to fly by. Because the service center does not have regulatory authority over districts, permission from district personnel must be obtained prior to conducting the interviews. Contact is made with the central office, usually with the supervisor, followed by a call to the campus principal. Teachers are then contacted. If they cannot be reached at the initial phone call they are asked to return the call. If possible, interviews are scheduled during the conference period. When interviews are scheduled more than a week in advance, a reminder postcard is sent. Interviews are scheduled at the convenience of the teachers. This does not always coincide with what is convenient for evaluation staff, resulting in a loss of time.

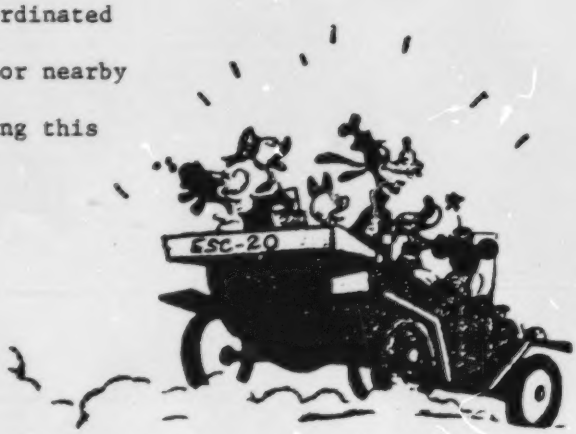
Conducting the Interview

On the day of the interview, the evaluator first goes to the principal's office to state that he or she is at the school. In some cases the principal may request the evaluator to check into the office before leaving the school. The interviewer then goes to the teacher's classroom or other designated site. The first part of the interview focuses on the criteria for the program. For each criteria, the interviewer attempts to identify how well a teacher is meeting the criteria. Once the information on the criteria is collected the interviewer proceeds to use the standard LoU format. (See Appendix 1) The interview is usally completed within thirty minutes.

At the conclusion of the interview, additional interview information may be collected. This information spans aspects of the program not covered by implementation criteria, such as quality of materials or the provision of technical assistance. Supplemental information also has been gathered success-

fully by other users (Klenke and Barrows, 1980).

Whenever possible, travel is coordinated so that several teachers in the same or nearby campuses are interviewed. Coordinating this activity to be cost effective has proved difficult.



Data Analysis

All interviews are transcribed. Essential for quality tapes are a top-of-the-line tape recorder and a secretary familiar with educational programs and their related jargon is a definite asset. Transcribing is done for two reasons. First, it is cost effective. A person can read faster than they can hear or write. The cost of transcribing is offset by evaluator time saved, especially when a summary of twenty interviews, each with five criteria, must be prepared. Quotes are included in evaluation summaries to enhance authenticity. Time is saved if the evaluator is not required to review the tape and transcribe information. Second, transcripts can be reviewed and the accuracy of data verified. The transcripts themselves are a good management tool. Evaluating programs using LoU interviews represents

a new activity for many Region 20 evaluation staff. Reviewing transcripts with a knowledgeable colleague can be a beneficial learning experience leading to a higher quality final report.

Interview information is analyzed in a two step process. The first step is to determine what percentage of the teachers are implementing based on whether or not they meet critical criteria. The summative objective called for 60% of the teachers to be implementers (Levels III, IV-A, IV-B, and V). The findings shown by Table 1 reveal that 65% were implementing with 45% of the teachers being at the lowest implementation level, Level III, mechanical.

Table 1

Level of Use Rating of READS Teachers
(N = 20)

Level	Percentage of Teachers
0 - Non use	35%
I - Orientation	0%
II - Preparation	0%
III - Mechanical	45%
IVA - Routine	15%
IVB - Refinement	0%
V - Integration	5%
VI - Renewal	0%

The second step is to analyze interview information according to both critical and related criteria. How each teacher implemented the innovation according to criteria is recorded. Similar patterns or dominant configurations emerge (Hall, Zigmari, & Hord, 1979). Analyzing data by configuration presently is not done in our setting, although future research efforts,

could profit by examining configurations. The relatively few teachers interviewed limited the number of potential dominant configurations, however.

The best information was obtained by analyzing across criteria. Staff found this information most relevant to immediate program needs. One of the related criteria, for example, required teachers to group students by ability level. All teachers met this criteria. Four variations were identified. The most popular grouping methods were completely individualized (40%) and a combination of large and small group techniques (30%). Figure 2 is a schematic depicting how READS data could be analyzed either by criteria (across) or configurations (down). In conclusion, data analyzed across criteria instead of by configuration best meets the unique curricular improvement demands of program staff.

Figure 2

Implementation for Reads Teachers

Grouping	Teacher Identification Number: 1 2 3 4 . . . . . 11 12 13 14 . . . . 19 20
Critical Criteria	
1. Student Objectives	
. . . . .	★★★★
. . . . .	★ ★
. . . . .	★ ★ ★
Related Criteria	
1. Identify Reading Level	
. . . . .	★ ★ ★
. . . . .	★★★★
. . . . .	★ ★
2. Grouping	
Completely individualized (one to one)	★★★★
Begin all students at one level or one of two levels, but then begin grouping students based on their performance.	★ ★
Combination of large and small group techniques.	★ ★ ★
3. Testing for Skill Mastery	
. . . . .	
. . . . .	★★★★ ★ ★ ★ ★ ★
4. Management System	
. . . . .	★ ★ ★
. . . . .	★ ★
. . . . .	★★★★

DOMINANT CONFIGURATIONS

DISCUSSION AND IMPLICATIONS

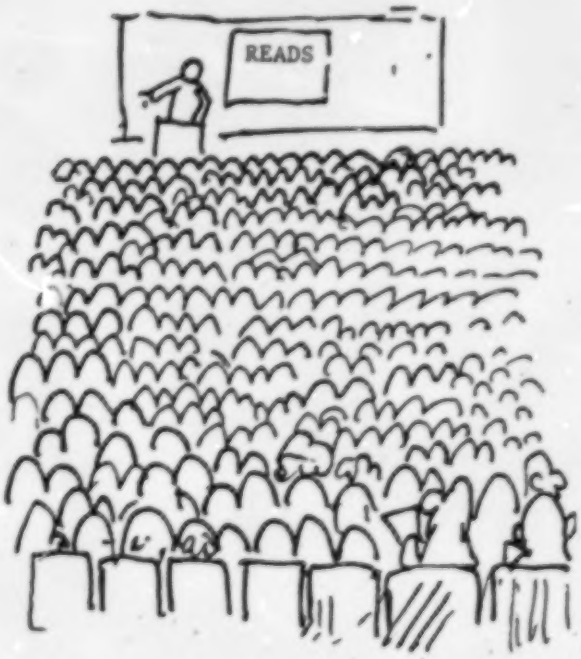
One of the few studies available on program implementation measured by LoU is similar to READS. This evaluation studies the reading component of an Individually Guided Instruction Program (IGE) in the Austin Independent School District. A general comparison reveals READS was not as well implemented as IGE. Thirty-five percent of READS teachers were non-users, compared to only 20% for IGE. More significant, most READS teachers were still struggling to get the program off the ground (LoU III). Most IGE teachers, like teachers from other studies (Hall & Loucks, 1977; Loucks, Hall, Rutherford, Newlove & George, 1976; Reidy & Hord, 1979) were implementing at the routine level (LoU IV-A). The somewhat regressed implementation shown by READS participants can be explained by the complexity of instructional design and certain incorrect assumptions about how implementation proceeds.

The assumption supporting many Center programs like READS was that most teachers trained would be implementing after one year. This was not valid. Subsequent experience with other Region 20 programs demonstrated that 40 to 60% of teachers were implementing after one year and up to 80% were implementing after two. The implementation process, stated Loucks et al (1976) takes considerable time.

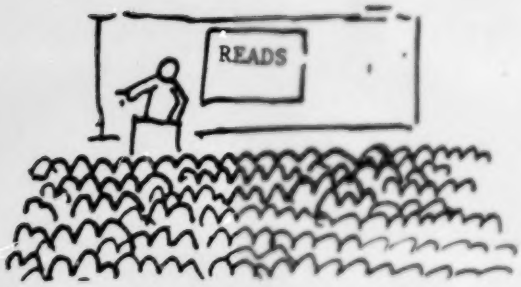
> ...our research documents the fact that implementing...takes <u>time.</u> It appears that teachers and professors, when implementing innovations, 'grow' developmentally in their concerns and use of the innovation. They are not simply non-users of an innovation one day and sophisticated users the next. (p. 12)

Programs built on the assumption that the implementation takes just one year were, by design, ineffectual. Such programs did not account for the large number of previous years teachers who were not implementers. As programs continued, more teachers needed to be serviced--usually with the same or a reduced level of program resources.

Some Years Later



First Year



The long-term result was that programs were less efficient and teachers did not receive the support required. Fewer teachers than expected became implementers; those who did implement, did so at a lower level. Such was the case with READS.

Simpler innovations are easier to implement than complex ones. Facilitators can install them quicker and teachers can understand them readily. Simplicity is the key, especially when one remembers that the agency supplying the seed money wants to see near-immediate results. The management system underlying READS was a complicated one, with similar systems frequently requiring computer assistance for implementation (Baker, 1978; Roecks, 1979b).

Instructionally, many of the same systems are not sound (Roecks, 1979a). One goal of management systems like READS is to save teacher time. For READS an excessive amount of teacher time was required, as illustrated from teacher comments:

> I finished charting, well, the student profile charts. I finished that last week and here it is December and I have three (READS) classes. If I would have had five, I still would be working on that (User).

> We feel very strongly that it is too complicated, too much paper work, and it takes away from actual classroom instruction (Non-user).

Strengths and Weaknesses

Region 20 program staff contend that LoU interviews provide some of the best curricular evaluation information available. They believe LoU interviews give them unique information for improving their programs. Staff are disappointed if the evaluation does not employ LoU interviews. They are no longer content with survey information or checklists.

At least four reasons can be cited why LoU interview information is useful for program improvement. First, information is provided that program staff cannot obtain on their own. Region 20 evaluation staff can secure higher quality information due to their organizational independence from the program. Participants are reluctant to "tell it like it is" to those they work with. Moreover, confidentiality of information is guaranteed by the policy of our evaluation office. Second, LoU information corroborates whether or not the program is proceeding according to plan. The fidelity of implementation is shown by the degree to which LoU findings correspond to critical and related criteria (See Figure 1). Third, the process of having to develop specific criteria for the program has potential side affects. Evaluation and program staffs develop a stronger working relationship and both gain an understanding of the program. The goal of the process, well defined relevant criteria, ultimately benefits the classroom teacher state Hall and Loucks (1978):

> ...it appears that when the developer is not clear in describing the innovations or change facilitators do not communicate effectively. users are apt to implement an unusual configuration or at least describe the innovation in different ways. On the other hand, if the implementation process is well designed and articulated, and is consistent with the developer's model, the semantic and operation confusion is not so apt to occur. (p. 30-31)

Fourth, the kind of information resulting from LoU interviews generally cannot be obtained from survey instruments. Questionnaire data represents one way communication where teachers are asked to respond to program and evaluation

staffs' preconceptions of how teachers are using the program. This "Catch 22" phenonema can be summarized as follows:

1. Questionnaires are sent out to determine what the teachers are doing with the innovation.
2. The only way the questionnaires will collect this information is if the questions asked are relevant to what the teachers are supposed to be doing.
3. There is no way to ensure that the questions are relevant without already knowing what the teachers are doing.
4. If you already know what the teachers are doing, there is no point in sending out the questionnaires.

Interviews are most effective in correctly identifying how teachers are using the innovation. While the interviewers have a predetermined list of criteria, they are free to go into other aspects or issues the teachers themselves raise. Unanticipated outcomes, which do not directly relate to implementation criteria, provide some of the most useful evaluation data. New criteria may emerge.[5] Variations occur in both the users and the innovation itself.

[5] Although evaluation in our setting is objective based, important side effects can be identified by using techniques such as LoU. Examining program effects irrespective of goals has been advocated by Michael Scriven (1974), a pioneer in the area of goal-free evaluation.

Illustrative findings include:

Student Population Served:

- READS was used almost exclusively with remedial or special education students.
- Some special education teachers were unaware that READS could be used in small group settings.

Resource Materials:

- Some middle school teachers have problems finding materials that their remedial students would not see as being "baby books."
- Teachers frequently did not have access to the materials listed for each objective.
- M st teachers wanted the READS cards laminated.

READS Objectives:

- About half the middle school teachers believed objectives were too specific.
- Some teachers perceived a conflict between READS objectives and state and district guidelines.

Testing and Class Management

- Some teachers believed the pretest placed students at too high a reading level.
- Several teachers found pretesting required months to complete.
- Posttesting by objectives was sometimes not done because teachers needed more than one set of the posttests and reproduction costs were high.
- Some teachers were having trouble managing behavioral problems in small group settings.

The biggest drawback to LoU interviews is the cost associated with securing information useful for program modification. Carrying out the interviews and determining the overall LoU is easy to accomplish. This task is inexpensive relative to the cost of giving formative information. The cost to carry out an evaluation study using LoU for 20 teachers is given by Table 2. About three person weeks, representing between 50% to 60% of the total cost, are required to complete the study. Transcribing costs are $20 to $25 per tape.

The cost given by Table 2 assumes personnel are trained in applying LoU as a formative evaluation tool. Evaluators must undergo a three day training session in order to be considered for certification. Additional time, reviewing pre-rated tapes, is required for certification. Learning how to compile LoU data for the purpose of formative evaluation is a time consuming activity.

Table 2

Cost to Carry Out an Evaluation Study Using LoU for 20 Teachers[a]

Expenditure	Cost Estimate	
	Low	High
Personnel (about three person weeks)[b]	$ 950	$ 1,250
Transcribing 20 tapes[c]	400	500
Overhead (travel, tape recorder, tapes, office equipment, supplies)[d]	350	550
Total	$ 1,700	$ 2,300

[a]Interview information is used for both summative and formative purposes. If only summative information is required (overall LoU), the cost is 30 to 40 percent lower.

[b]Personnel should be certified in Levels of Use. Cost includes loss of three workdays plus cost of training and travel expenses. These costs are not included.

[c]This represents the amount paid to an outside business. We do our own transcribing, resulting in higher quality at a slightly lower cost. Initial investment includes a transcribing unit ($850.00) and typewriter ($800.00).

[d]As Region 20 serves 50 schools in fourteen counties much of the overhead cost is travel. Tape recorders (Lanier Dictaphones) cost $250.00, tapes $4.00.

Summary

This study presented an evaluation of a reading program whose intent was to train over 200 K-8 teachers over a three year period. A sample of 20 teachers was interviewed using the Levels of Use technique. Criteria for successful implementation were identified. Findings revealed 65% of the teachers implementing at various levels, with the degree of implementation being lower than for similar studies. The regressed implementation was accounted for by the complexity of instructional design and the incorrect assumption that most teachers would be implementing after one year.

Of further significance, was the wealth of formative evaluation information provided by interviews in the areas of student characteristics, materials, objectives, testing and classroom management. This additional information identified common practices and the most frequent ways that teachers implemented the program, prevailing problems, and the typical organizational patterns supporting the innovation within the school. The information proved useful in identifying areas of need and allowed project staff to make data-based decisions targeted at program improvement. The employment of Levels of Use interview information was especially valuable in its application to the purposes of formative evaluation by practitioners responsible for effective change in schools. The major limitation cited for the technique was the cost to carry out valid evaluation.

References

Anderson, S., & Ball, S. The profession and practice of program evaluation. San Francisco: Jossey-Bass, 1978.

Baker, F. B. Computer managed instruction. Englewood Cliffs, New Jersey: Educational Technology Publications, 1978.

Berman, P., & McLaughlin, M. W. Federal programs supporting educational changes, Vol. VIII: Implementing and sustaining innovations. Santa Monica, California: Rand Corporation, 1978.

Charters, W. W., Jr., & Jones, J. E. On the risks of appraising non-events in program evaluation. Educational Researcher, 1973, 2(11), 5-7.

Drezek, S. F., Estes, D. W., Roecks, A. L., & Andrews, J. H. Program objectives. Paper presented at the annual meeting of the Southwest Educational Research Association, San Antonio, February 1980.

Freeman, H. D. The present status of evaluation research. In M. Guttentag (Ed.), Evaluation studies review annual. Beverly Hills: Sage Publications, 1977.

Fullan, M., & Pomfret, A. Research on curriculum and instruction implementation. Review of Educational Research, 1977, 47(1), 335-393.

Gebhardt, M. E. On the social and political barriers to effective educational training and evaluation. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April 1979.

Hall, G. E., & Loucks, S. F. A developmental model for determining whether the treatment is actually implemented. American Educational Research Journal, 1977, 14(3), 263-276.

Hall, G. E., & Loucks, S. F. Innovation configurations: Analyzing the adaptations of innovations. Paper presented at the annual meeting of the American Educational Research Association, Toronto, March 1978.

Hall, G. E., & Loucks, S. F. Teacher concerns as a basis for facilitating and personalizing staff development. Teachers College Record, 1978, 80(1), 36-53.

Hall, G. E., Loucks, S. F., Rutherford, W. L., & Newlove, B. W. Levels of use of the innovation: A framework for analyzing innovation adoption. Journal of Teacher Education, 1975, 26(1), 52-56.

Hall, G. D., & Rutherford, W. L. Concerns of teachers about implementing team teaching. Educational Leadership, 1976, 34(3), 227-233.

Hall, G. E., Zigarmi, P. K., & Hord, S. M. A taxonomy of interventions: The prototype and initial testing. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April 1979.

Klenke, W. H., & Barrows, L. K. Documenting change: Procedures, problems, and possibilities. Paper presented at the annual meeting of the American Educational Research Association, Boston, April 1980.

Loucks, S. F., Hall, G. E., Rutherford, W. L., Newlove, B. W., & George, A. The concerns-based adoption model applied to the ISCS science curriculum: Research report to participants. Austin, Texas: Research and Development Center for Teacher Education, The University of Texas, 1976.

Mushkin, S. J. Evaluations: Use with caution. Evaluation, 1973, 1(2), 30-35.

Okpalobi, M. J. An exploration of levels of use of the innovation in the evaluation of a teaching reading in content areas staff development project. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April 1979.

Reidy, E. F., & Hord, S. M. Utilizing implementation data in the evaluation of a mathematics skills achievement monitoring program. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April 1979.

Rist, R. C. On the relations among educational research paradigms. Anthropology and Education, May 1977, pp. 42-45.

Rossi, P. H. Issues in the evaluation of human services delivery. Evaluation Quarterly, 1978, 2(4), 573-599.

Rossi, P. H., Freeman, H. E., & Wright, S. R. Evaluation: A systematic approach. Beverly Hills, California: SAGE, 1979.

Roecks, A. L. The MICA project: A final look. AEDS Journal, 1979, 12(3), 146-154. (a)

Roecks, A. L. Securing local support for your computer project. Educational Technology, 1979, 19(7), 16-18. (b)

Scriven, M. Pros and cons about goal-free evaluation. In W. J. Popham (Ed.), Evaluation in education. Berkeley, Calif.: McCutchan Publishing Corporation, 1974.

Stufflebeam, D. L., Foley, W. J., Gephart, W. J., Guba, E. G., Hammond, R. L., Merriman, H. O., & Provus, M. M. Educational evaluation & decision making. Bloomington, Indiana: Phi Delta Kappa, Inc., 1971.

Weiss, C. H. Utilization of evaluation: Toward comparative study. In C. H. Weiss (Ed.), Evaluating action programs. Boston: Allyn and Bacon, 1972.

This paper was one of four in the 1980 AERA Symposium on "Documenting Educational Change: Practical and Research Applications of the Levels of Use Interview." Other papers and participants include:

*George, Archie and Hord, Shirley, Research and Development Center for Teacher Education, University of Texas, Austin, Texas 78712
"Monitoring curriculum implementation: Mapping teacher behaviors on a configuration continuum"

Hall, Gene (discussant) and Heck, Susan (chairperson), Research and Development Center for Teacher Education, University of Texas, Austin, Texas 78712

*Klenke, William and Barrows, Linda, Wisconsin Research and Development Center, University of Wisconsin, 1025 West Johnson, Madison, Wisconsin 53706
"Documenting Change: Procedures, Problems, and Possibilities"

*Knowles, Claudia, Department of Physical and Health Education, University of Texas, Austin, Texas 78712
"Levels of Use of Individualized Instruction in the Physical Education Setting"

Roecks, Alan and Andrews, John, Education Service Center, Region 20, 1550 N. E. Loop 410, San Antonio, Texas 78209

Tabachnick, Robert (discussant), Department of Curriculum Instruction, University of Wisconsin, Madison, Wisconsin 53706

LoU Interview

0-II/III-VI	Are you currently using _____?

NO

Have you ever used it in the past? If so, when? Why did you stop?
(if yes, go to *, then return)

0/I-II	Have you made a decision to use _____ in the future?
I/II	If so, when will you begin use?
Knowledge	Can you describe _____ for me as you see it?
Acquiring Information	Are you currently looking for any information about _____? What kinds? For what purposes?
Knowledge	What do you see as the strengths and weaknesses of _____ in your situat'on?
Assessing	At this point in time, what kinds of questions are you asking about _____? Give examples if necessary.
Sharing	Do you ever talk with others and share information about _____? What do you share?
Planning	What are you planning with respect to _____? Can you tell me about any preparation or plans you have been making for the use of _____?
Final Question (Optional)	Can you summarize for me where you see yourself right now in relation to the use of _____?

PAST USERS*

Can you describe for me how you organized your use of _____, what problems you found, what its effects appeared to be on students?

When you assess _____ at this point in time, what do you see as the strengths and weaknesses?

(Return to other nonuse questions.)

YES

Open-ended	Please describe for me how you use _____. (Ask sufficient questions to get configurations.)
Assessing/ Knowledge	What do you see as the strengths and weaknesses of _____ in your situation? (Have you made any attempt to do anything about weaknesses? Probe those they mentioned specifically.)
Acquiring Information	Are you currently looking for any information about _____? What kind? For what purposes?
LoU V	Do you work with others in your use of _____? Do you meet on a regular basis? Have you made any changes in your use of _____ based on this coordination? (if yes, go to *)
→Sharing	Do you ever talk with others about _____? What do you tell them?
Assessing	(Have you considered any alternatives or different ways of doing things with the program?) Are you doing any evaluating, either formally or informally, that would affect your use of _____? Have you received any feedback from students that would affect the way you're using _____? What have you done with the information you got?
III/IVA/IVB	Have you made any changes recently in how you use _____? What? Why? How recently? Are you considering making any changes?
Planning/ Status Reporting	As you look ahead to later this year, what plans do you have in relation to your use of _____?
→III-V/VI	Are you considering or planning to make major modifications or replace _____ at this time?

*LoU V Probes

1. How do you work together? What things do you share with each other?
2. How frequently?
3. What do you see as the effects of this collaboration?
4. Are you looking for any particular kind of information in relation to this collaboration?
5. Do you talk with others about your collaboration? If so, what do you share with them?
6. Have you done any formal or informal evaluation of how your collaboration is working?
7. What plans do you have for this effort in the futur ?

If you have enough evidence to place the person at an LoU V, go to ⌐

If you do not think the person is an LoU V, go to ⌐

Happy Endings

WHY LoU INTERVIEWS WORK.

1. UNIQUE PRODUCT
   - EVALUATION INDEPENDENCE
   - CONFIDENTIALITY
   - CAN'T GET FROM SURVEYS
2. DOCUMENTS PROGRAM STATUS
3. IMPLEMENTATION CRITERIA → BETTER DEFINITION
4. UNANTICIPATED OUTCOMES

I HAVE TO BE FED TO
DO QUALITY
WORK!!!